

**ACM CIKM 2022**

***International Workshop on Privacy Algorithms in Systems***

# **Privacy in the Age of AI and the Internet of Things**

*Norman Sadeh*

Professor of Computer Science

Co-Director, Privacy Engineering Program

Carnegie Mellon University

<https://normsadeh.org>

# Privacy Threats Are Everywhere



Source: CSO online

## Data-Hungry Economy

- AI/ML
- IoT sensors everywhere
- Myriads of APIs and dataflows and also:
- unscrupulous/ignorant data collectors/processors

# New Regulations

- Increasingly more specific data practice disclosure requirements
- Increasingly specific data subject rights
- Emerging, yet loosely specified, usability expectations
- AI & Privacy – Broadening Expectations (e.g., Interpretability)
- New, significantly steeper financial penalties



# *“All is for the Best in the Best of Possible Worlds” (Voltaire)*

Privacy Policy Before GDPR			
COMPANY	WORD COUNT	READING TIME *	READING LEVEL
Google	2722	12:54	14
Facebook	5420	26:24	11
Reddit	5524	26:42	12
Amazon	2627	12:48	12
Wikipedia	2881	13:49	15
Yahoo	1611	7:44	14
Twitter	3764	18:37	13
eBay	5244	25:27	18
Instagram	2981	14:26	12
Netflix	3046	15:10	15

\* Minutes:Seconds



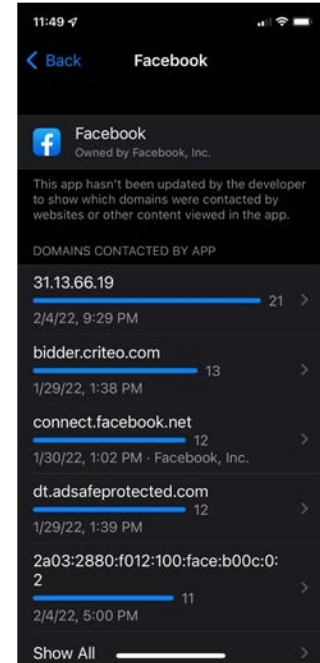
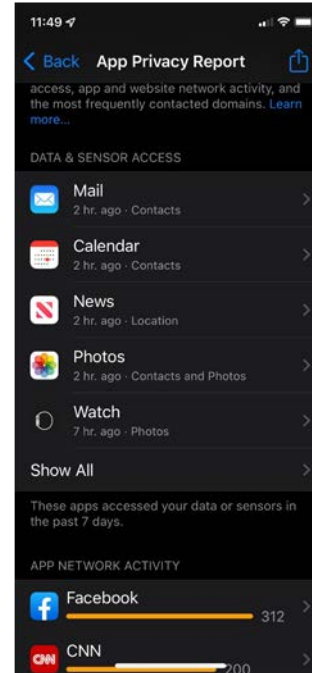
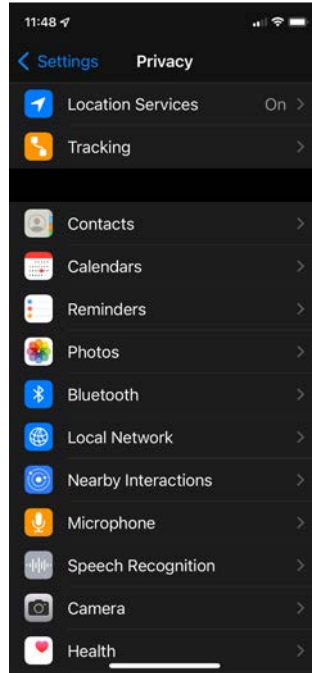
Progress?

Privacy Policy After GDPR			
COMPANY	WORD COUNT	READING TIME *	READING LEVEL
Google	4036	19:11	14
Facebook	4233	20:41	13
Reddit	3414	16:39	12
Amazon	3837	18:24	13
Wikipedia	5617	27:06	14
Yahoo	2225	11:12	13
Twitter	4880	22:25	13
eBay	5666	27:32	20
Instagram	4221	20:38	13
Netflix	3417	16:39	16

\* Minutes:Seconds

Source: <https://www.varonis.com/blog/gdpr-privacy-policy>

# ...And Who Has the Time to Review & Manage All these Privacy Settings?



# The Human Bottleneck

Lack of:

- **Expertise**
- **Time**
- **Attention**
- **Motivation**
- **etc.**



Source: <https://www.datanami.com/2016/09/13/sas-goes-back-future-cognitive-computing-viya/>

# What If Computers Understood the Text of Privacy Policies?

# Annotation Tool

Current Policy: a.98\_neworleansonline.com

**Select a category**

First Party Collection/Use | Third Party Sharing/Collection

User Choice/Control | User Access, Edit and Deletion

Data Retention | Data Security | Policy Change | Do Not Track

International and Specific Audiences | Other

7/41

Previous | Next

Annotated Practices: 1

**Select an attribute**

**Select a value**

**Highlight text span for an attribute, value pair**

**Information We Collect**

Whether you access our Online Services from your computer, smart phone, tablet or other mobile device, NOTMC and its agents may collect some information that identifies you or relates to you as an individual ("Personal Information"), such as your name, email address, telephone number, credit address, user name and password (for account administration), device ID, including IP address, geolocation (if using a mobile application and you consent to providing it), and additional personal information necessary for the administration of certain promotional events.

Please write your comments for this paragraph:

**Practices of this paragraph**

First Party Collection/Use

- Does Unspecified Collect on website Identifiable Contact Unspecified Unspecified Unspecified Unspecified [Clone](#) [Delete](#)

Third Party Sharing/Collection

**First Party Collection/Use**

Does/Does Not: Does

Collection Mode: Unspecified

Action First-Party: Collect on website

Identifiability: Identifiable

Personal Information Type: Contact

Purpose: Unspecified

User Type: Unspecified

Choice Type: Unspecified

Choice Scope: Unspecified

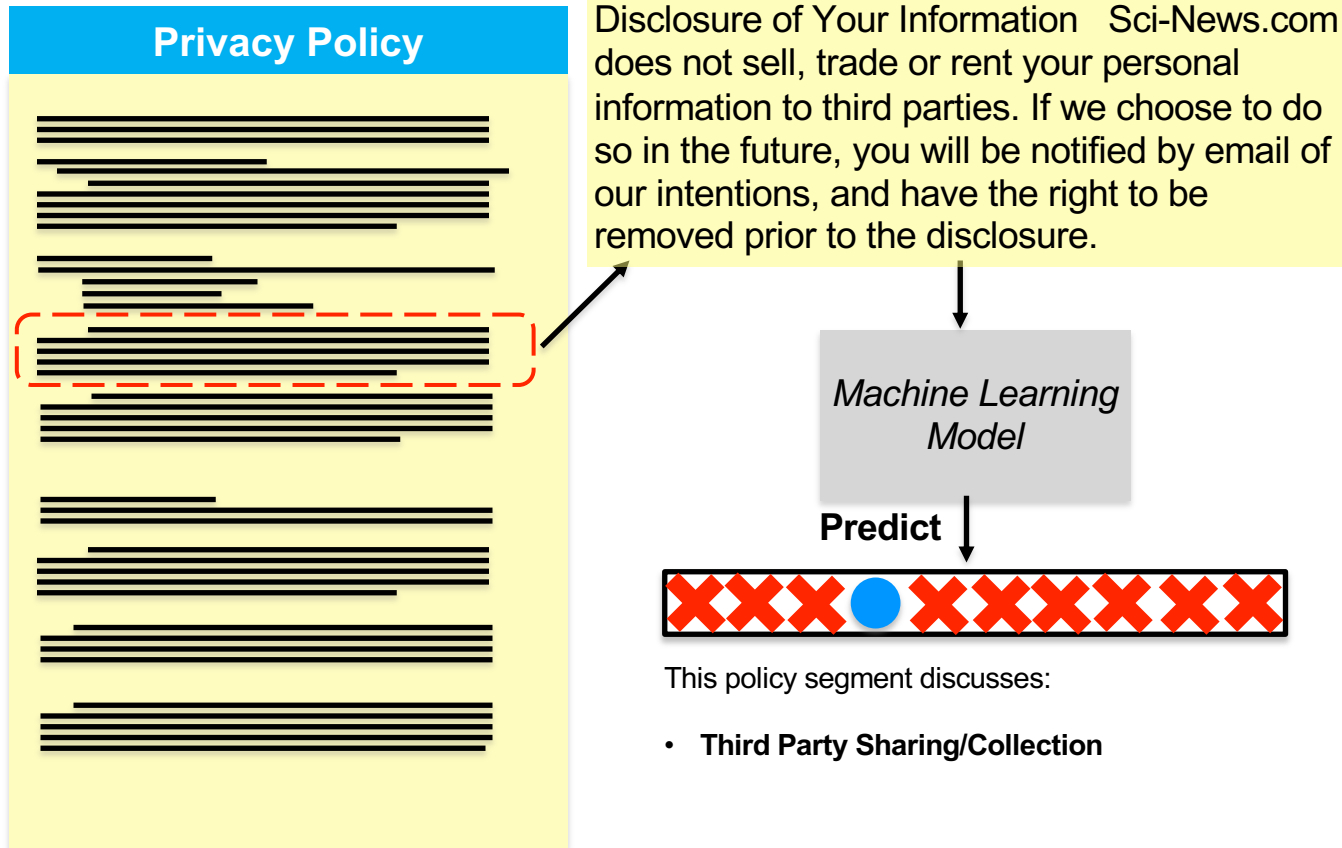
☐ References another place in the policy

[Save](#)

S. Wilson, F. Schaub, A. Dara, F. Liu, S. Cherivirala, P.G. Leon, M.S. Andersen, S. Zimmeck, K. Sathyendra, N.C. Russell, T.B. Norton, E. Hovy, J.R. Reidenberg, N. Sadeh, "The Creation and Analysis of a Website Privacy Policy Corpus", ACL '16: Annual Meeting of the Association for Computational Linguistics, Aug 2016



# A First Task: Segment Annotation



# Automatic Identification of Data Practice Disclosures

Yahoo! yahoo.com

Arts Business Computers Games Health Home Recreation Reference Regional Society World

Take a tour

## Privacy Practices

Click a category to filter practice statements.

First Party Collection/Use 67

Third Party Sharing/Collection 21

User Choice/Control 6

User Access, Edit and Deletion 3

Data Retention 1

Retention period

☒ All  
☐ Indefinitely (1)

Purpose of retention

☒ All  
☐ Unspecified (1)

more filters

Data Security 8

Policy Change 6

Do Not Track 0

International and Specific Audiences 5

## Privacy Policy

Yahoo News Privacy Policy from Sep 25, 2014.

Reading Level: College (Grade 13)

125 privacy practice statements in total

This privacy policy also applies to Flickr, Yahoo Finance, Yahoo News, Yahoo Sports, and Yahoo! Good Morning America.

We reserve the right to send you certain communications relating to the Yahoo service, such as service announcements, administrative messages and the Yahoo Newsletter, that are considered part of your Yahoo account, without offering you the opportunity to opt out of receiving them.

You can delete your [Yahoo account by visiting our Account Deletion page](#). [Please click here to read about information that might possibly remain in our archived records after your account has been deleted.](#)

### CONFIDENTIALITY

We limit access to personal information to those who we believe reasonably need to come into contact with that information to provide products or services to you or in order to do their jobs.

We have physical, electronic, and procedural safeguards that comply with federal regulations to protect personal information about you.

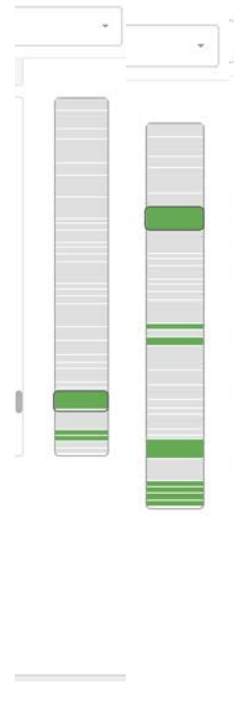
To learn more about security, including the security steps we have taken and security steps you can take, please read Security at Yahoo.

### CHANGES TO THIS PRIVACY POLICY

Yahoo may update this policy. We will notify you about significant changes in the way we treat personal information by sending a notice to the primary email address specified in your Yahoo account or by placing a prominent notice on our site.

### QUESTION AND SUGGESTIONS

If you have questions, suggestions, or wish to make a complaint, please complete a feedback



https://explore.usableprivacy.org/browse/category/

**USABLEPRIVACY.ORG** EXPLORE About Browse Privacy Policies Search for a website

## Browse

by **Category** Readability Popularity

- Arts 68
- Business 53
- Computers 42
- Games 26
- Health 35
- Home 37
- Kids and Teens 46
- News 32
- Recreation 42
- Reference 31

### Arts 68

**E! Online**  
Privacy policy from Jan 14, 2015 with 256 practice statements.

**FOX Sports**  
Privacy policy from Jun 11, 2015 with 215 practice statements.

**Racked**  
Privacy policy from May 1, 2014 with 204 practice statements.

See more

### Business 53

**Blogger**  
Privacy policy from Jun 30, 2015 with 241 practice statements.

**AOL**  
Privacy policy from Jun 23, 2015 with 232 practice statements.

**Allstate**  
Privacy policy from May 29, 2015 with 226 practice statements.

See more

# Press Coverage – Notice the Irony



FastCompany  
informing their readers  
about their new policy



FastCompany's article  
about our research



03.19.18

## You're Never Going To Read That Privacy Policy. Could AI Help?

This AI trained on legalese acts like a personal translator of confusing, opaque privacy statements.



ADVERTISEMENT



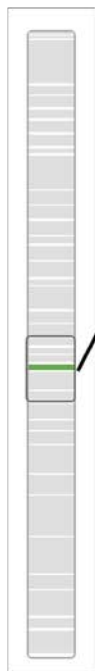
# Privacy Question Answering

- **One-size-fits-all summaries of privacy policies only go so far**
- Different people have different questions at different points in time
- Could we develop privacy question answering functionality?
- A number of challenges
  - Can people accurately articulate their questions. If not, how can we help them?
  - How do we provide useful answers – vague policies, inaccurate classifiers
  - etc.

Question answering for privacy policies: Combining computational and legal perspectives. A Ravichander, AW Black, S Wilson, T Norton, N Sadeh, EMNL 2019 Conference, arXiv preprint arXiv:1911.00841

Breaking Down Walls of Text: How Can NLP Benefit Consumer Privacy?, A Ravichander, AW Black, T Norton, S Wilson, N Sadeh, ACL/IJCNLP 2021. <http://dx.doi.org/10.18653/v1/2021.acl-long.319>

# User Choice Instance Extraction



**Choice Instance !!!**  
If you do not want us to use  
personal information that we  
gather to allow third parties to  
personalize advertisements  
we display to you, please  
adjust your Advertising  
Preferences .

- User choices often buried deep in the text of long policies
- Is it possible to **automatically extract information** about such “choice instances” from privacy policies?
- Use Natural Language Toolkit tokenizer to subdivide segments into sentences & build classifiers

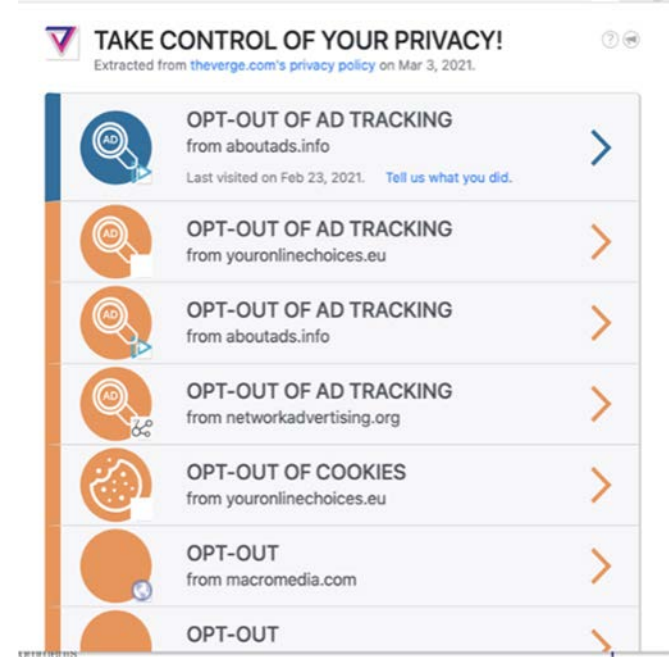
**Results: Recall & Accuracy > 90%**

Vinayshekhar Bannihatti Kumar, Roger Iyengar, Namita Nisal, Yuanyuan Feng, Hana Habib, Peter Story, Sushain Cherivirala, Margaret Hagan, Lorrie Faith Cranor, Shomir Wilson, Florian Schaub, Norman Sadeh, "Finding a Choice in a Haystack: Automatic Extraction of Opt-Out Statements from Privacy Policy Text", WWW '20, Apr 2020 [\[pdf\]](#)

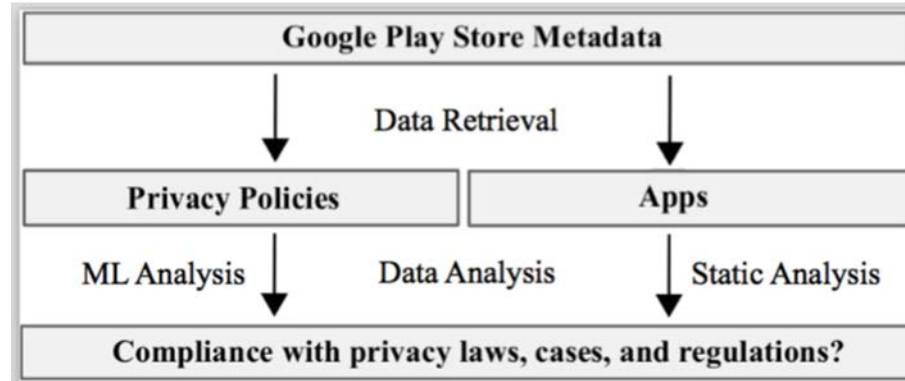
# Opt-Out Easy Browser Extension

- Automatically identify and categorize opt-out choices in the text of privacy policies
- And present them in an easy-to-use interface to users as they browse the web
- Available in Google Chrome store and Firefox store -

[Watch our video](#)



# Can We Automatically Check for Potential Compliance Issues?



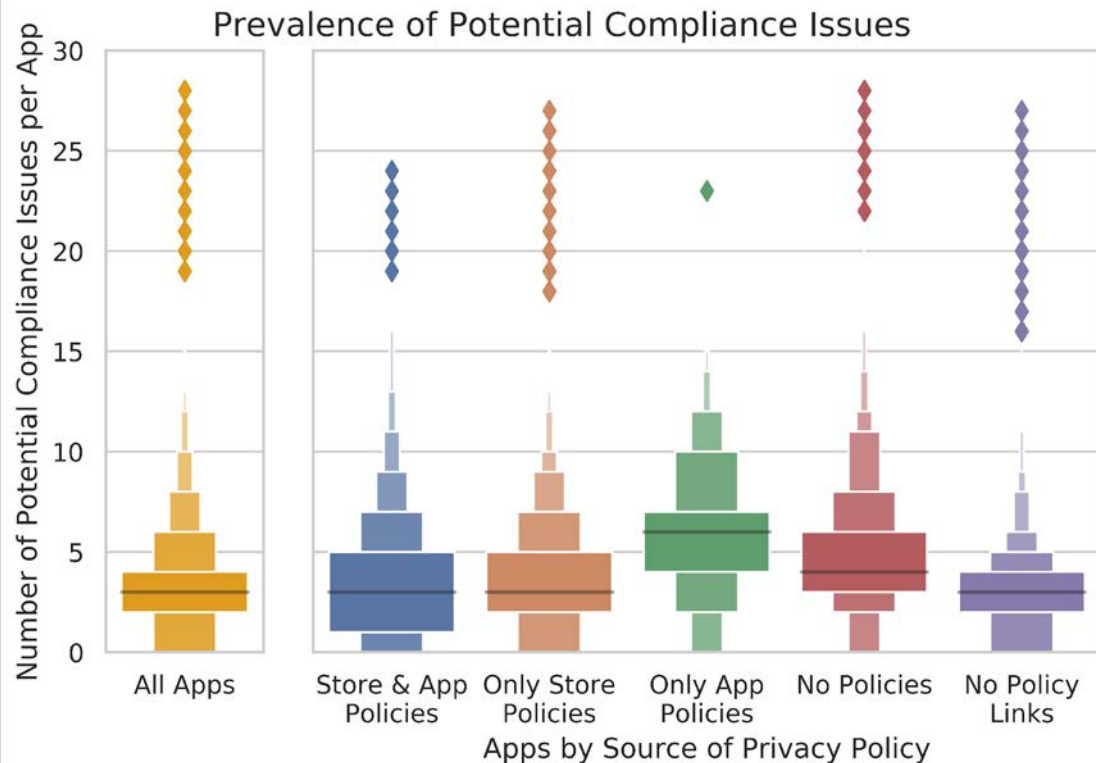
- Training **machine learning classifiers** to extract relevant policy statements
- Compare these statements against:
  - **Regulatory requirements**
  - What the software actually does
    - **Static and dynamic code analysis**

Zimmeck, S., Story, P., Smullen, D., Ravichander, A., Wang, Z., Reidenberg, J., Russell, N.C., Sadeh, N., "MAPS: Scaling Privacy Compliance Analysis to a Million Apps," in Proceedings on Privacy Enhancing Technologies, Vol. 3, pp. 66-86, 2019.  
<https://doi.org/10.2478/popets-2019-0037>

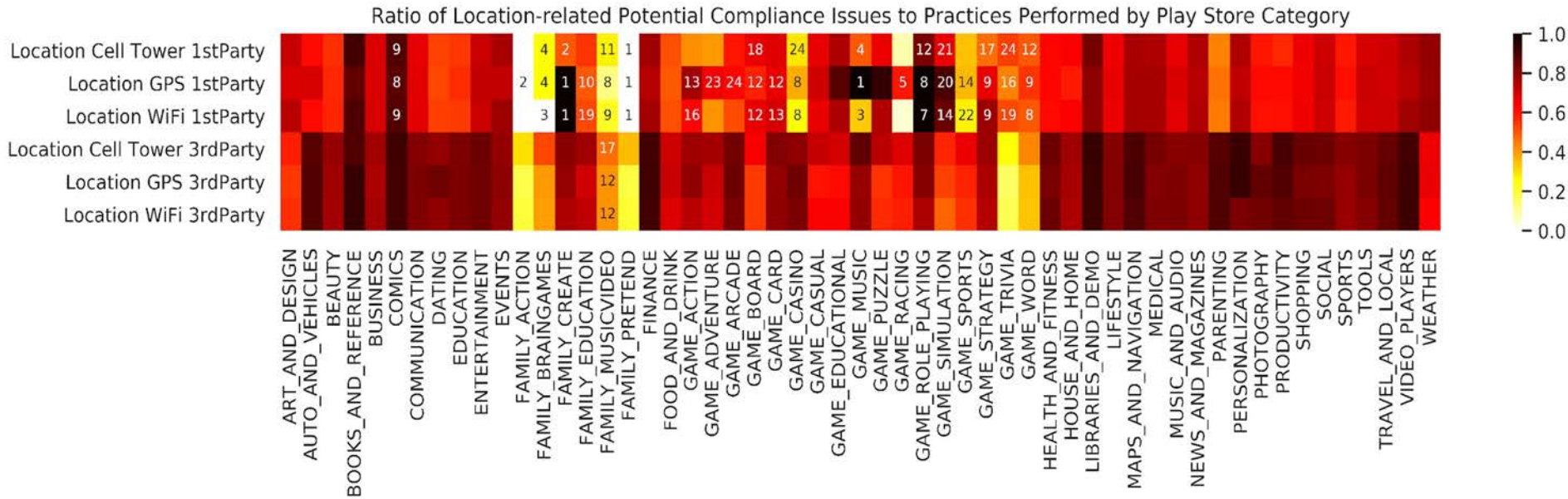


# Analysis of over 1 million Android Apps in Google Play Store

- Average number of potential compliance issues per app is 3.47 and the median is 3
- Requires manual vetting – both policy and app behavior to confirm potential compliance issue



# Developers Struggle with 3<sup>rd</sup> Party APIs



- Lighter colors indicate greater transparency of practices. Darker colors indicate that practices are being performed but not disclosed.
- Cells with fewer than 25 apps performing the practice are annotated with the respective number of apps.

# Other Collaborations

- Collaboration with California Attorney General's office
- COPPA report compiled for Federal Trade Commission
  - Focusing on location, apps with a large number of downloads, and companies based in the US
- CDT report on mobile apps for connected cars
- Work with large European electronics manufacturer – checking for GDPR compliance of mobile apps

# Tools for Developers



## 1. Template Provisioning

E.g., the GDPR requires policies to notify users of their rights to request data access, rectification, erasure, restriction of processing, objection of processing, and portability (Art. 13(2)(b))

Information to be provided where personal data are collected from the data subject  
Article 13  
1. Where personal data relating to a data subject are collected from the data subject, the controller shall, at the time when personal data are obtained, provide the data subject with all of the following information:  
(a) the identity and the contact details of the controller and, where applicable, of the controller's representative;  
(b) the contact details of the data protection officer, where applicable;

Xcode File Edit View Find Navigate Editor Product Debug Source Control Window Help

### CAMERA: USED

The use of camera information was determined because of these specific lines in your app's code:

Show Detected API and Third Party Library Calls

Code	<key>NSCameraUsageDescription</key>
File	missing-reports/krypton-ios-master/Krypton/Info.plist
Line Number	77
Used	USED
Usage Description	

You can customize the recommended statement about the detected data practice using the checkboxes below.

#### Specific Practices

- ☒ accesses camera data on user's device
- ☐ sends camera data to the developer's server
- ☐ stores camera data on user's device

#### Purposes

- ☒ performs a functionality of the app
- ☐ advertising
- ☐ analytics

You can adjust the recommended statement below and copy it into your application's privacy policy.

The app accesses camera data on your device for the purposes of the app's camera functionality [PLEASE DESCRIBE THE FUNCTIONALITY].

## 2. Code Analysis

Plist Permissions and Swift API Usage

```
1. Camera Information
2. Location Information
3. Photo Library Information
4. Address Book Information
5. Reminders Information
6. Calendar Information
7. Contacts Information
8. Microphone Information
9. Bluetooth Information
10. Motion & Core Location Information
11. Background App Refresh Information
12. Notification Information
13. Siri Suggestions Information
14. App Store Product Link Information
15. App Store Review Information
16. App Store Analytics Information
17. App Store Search Analytics Information
18. App Store Analytics Information
19. App Store Analytics Information
20. App Store Analytics Information
```

## 3. Wizard Fine Tuning

Don't worry! I got you covered. Here is the privacy analysis of your app.


You can adjust the generated privacy policy via the checkboxes.



Compliance traceability: Privacy policies as software development artifacts, S Zimmeck, P Story, R Goldstein, D Baraka, S Li, Y Feng, N Sadeh, Privacy, Usability and Transparency Workshop (PUT 2019) at PoPETs 2019 conference

# Could Computers Also Help Motivate People to Take Advantage of Privacy Settings?

# Nudging Users: Surprise People with Something That Will Motivate Them to Pay Attention

 Your location shared with 10 apps


Did you know?  
Your **location** has been shared **5398** times with Facebook, Groupon, GO Launcher EX, and 7 other apps for the past **14** days.

Let me change my settings






Show me more before I make changes

Keep sharing my location

Notification provided by AppOps.


 Your location shared with 10 apps

Number of times your **location** has been shared with each app for the past 14 days.






	Google Play services	1603
	Android System	1602
	Groupon	1602
	Weather & Clock Widget	296
	GO Launcher EX	255

Let me change my settings

keep sharing my location

 Your location shared with 10 apps

Number of times your **location** has been shared with each app for the past 14 days.

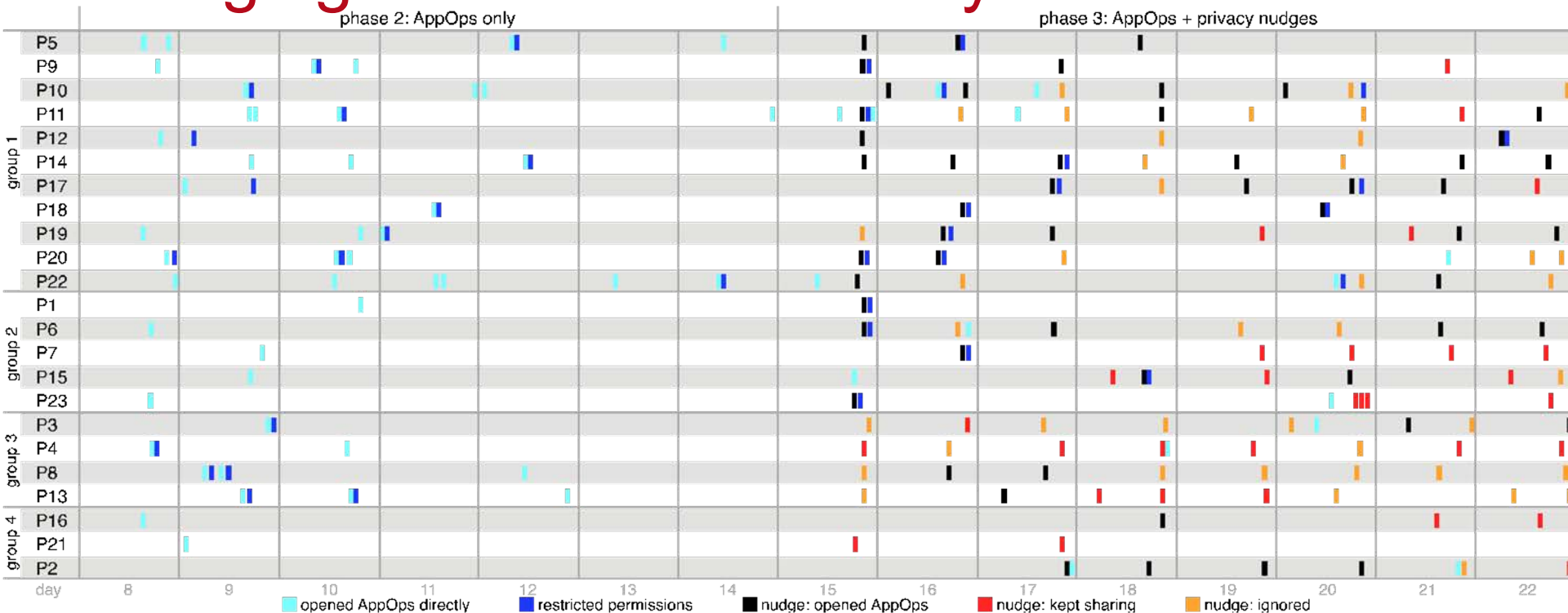
	Maps	18
	Viber	11
	Facebook	5
	Google Search	3
	MyFoodCoach Study	3

Let me change my settings

keep sharing my location

H. Almuhiemedi, F. Schaub, N. Sadeh, I. Adjerd, A. Acquisti, J. Gluck, L. Cranor, Y. Agarwal, "Your Location Has Been Shared 5,398 Times!: A Field Study On Mobile App Privacy Nudging," ACM CHI 2015.

# Nudging Pilot – 3-week study



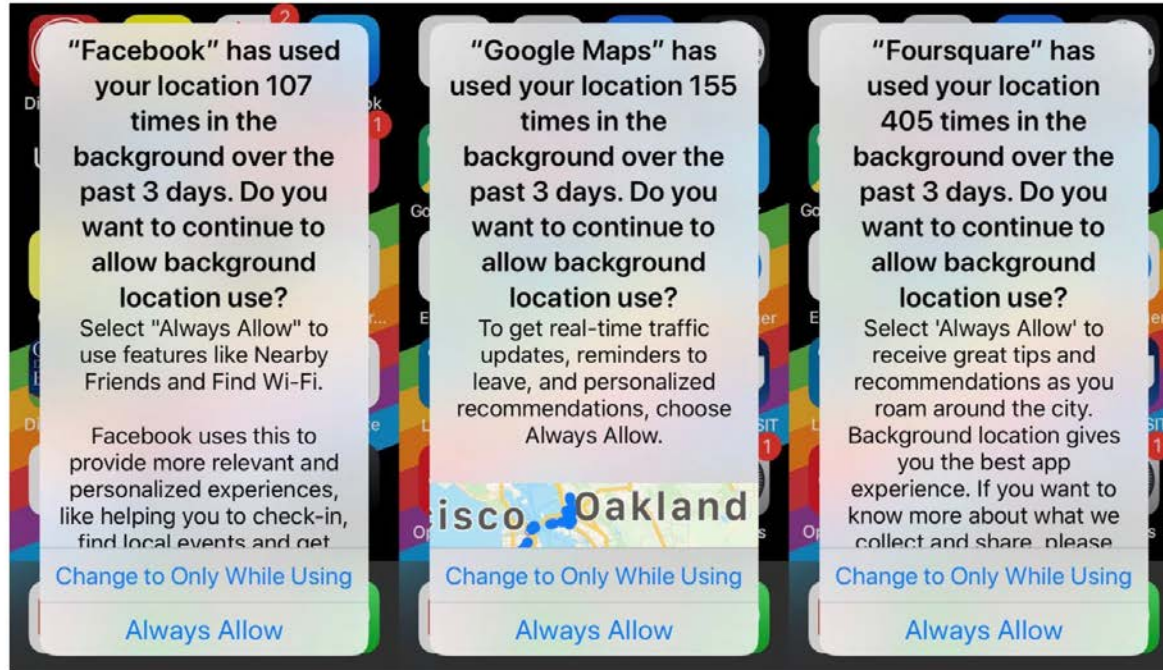
## Week 2: Permission Manager Only

## Week 3: Daily Nudges

- Permission Manager Helps (week 2)
- Nudges can make a big difference (week 3)



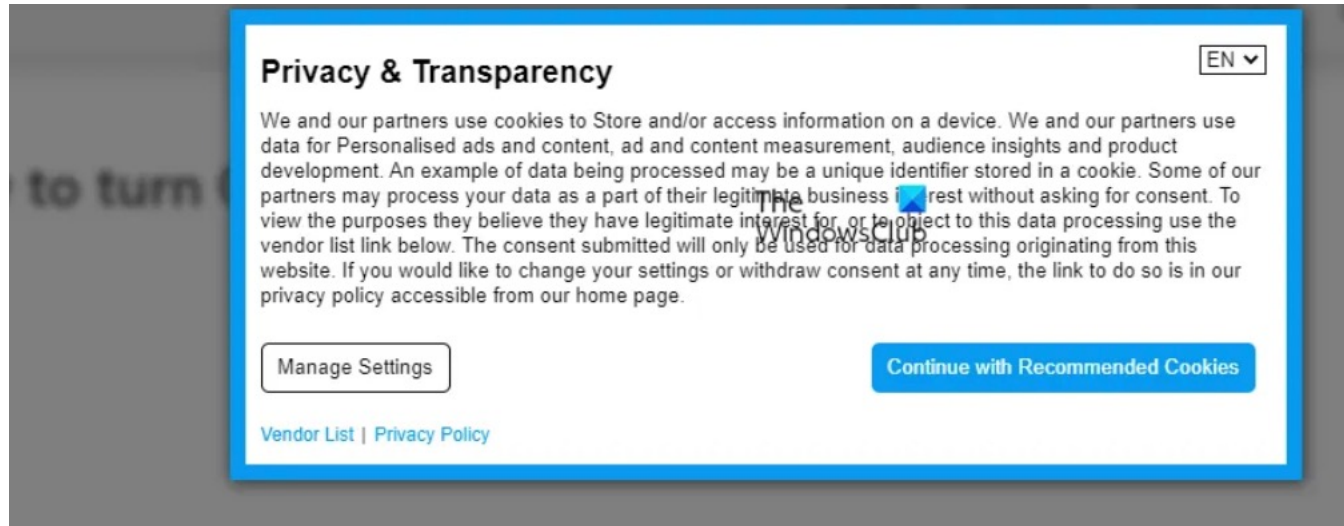
# Introduced in iOS13 – Privacy Reminders





# Could Computers Help People Manage their Privacy Controls?

# Many Privacy Decisions Are Repetitive, Similar and Tedious

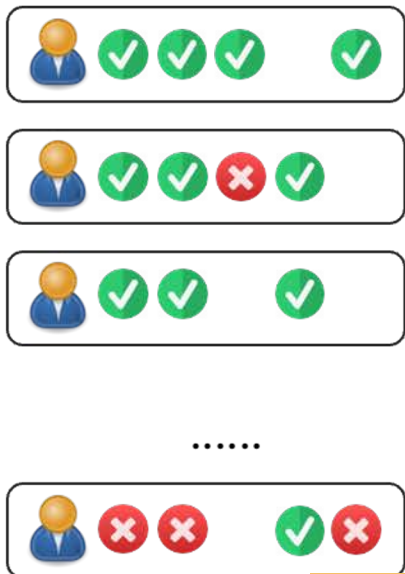


Standardized APIs could really help...

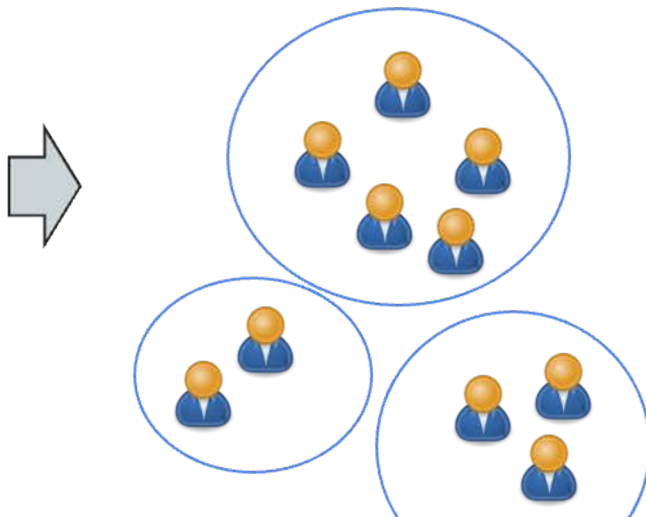
**Privacy Assistant:** Help users configure their privacy settings – motivates users, mitigate dark patterns, reduce user burden

# Privacy Assistants - I

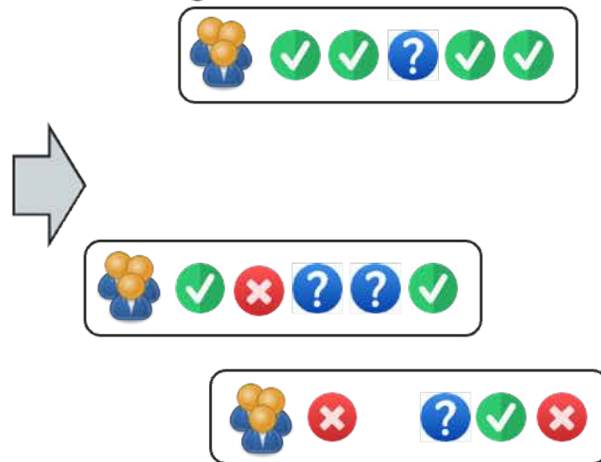
Users with their settings



Clustering of users based on features extracted from their settings



Each cluster has an associated set of recommended privacy settings



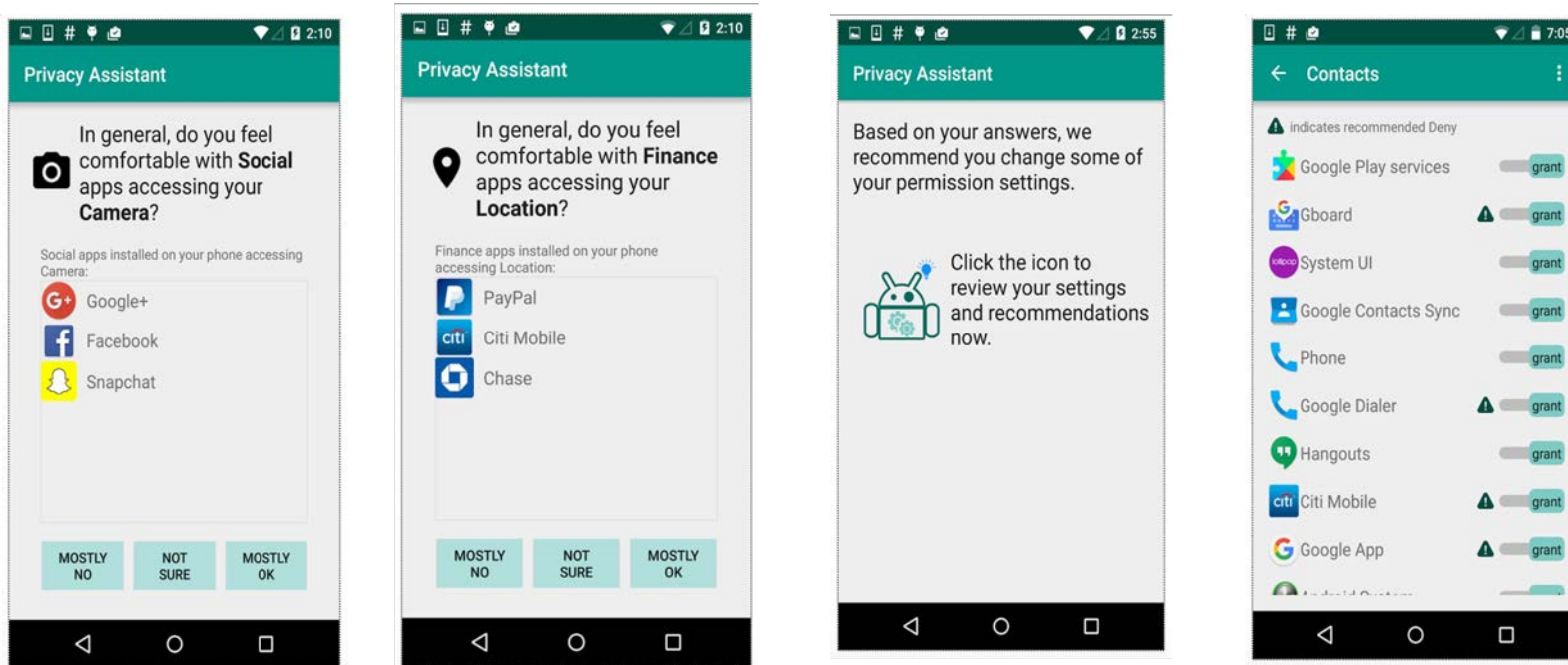
**Even simple solution with small number of clusters achieves high levels of accuracy**

B. Liu, M.S. Andersen, F. Schaub, H. Almuhiemedi, S. Zhang, N. Sadeh, A. Acquisti, and Y. Agarwal, "Follow My Recommendations: A Personalized Assistant for Mobile App Permissions", SOUPS 2016- US patents 10,956,586

© Copyright 2022, Norman Sadeh, All rights reserved.

# Privacy Assistants II

Generating recommendations rather than automating privacy decisions



**Vast majority of recommendations accepted by users and kept despite nudges to reconsider  
Successfully deployed in Google Play store for several users – rooted Android phones only...**

# Why Recommendations?

**Agency is a major part of privacy:** users should remain in charge of their decisions...but **AI can help** them make these decisions and can help overcome fundamental **usability limitations**

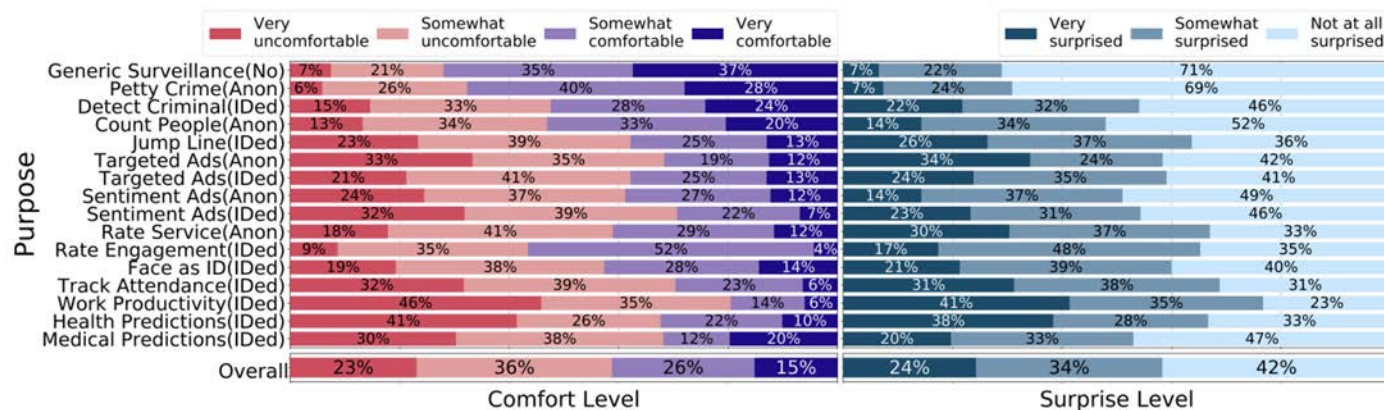
- **Major requirement:** the recommendations have to be ***understandable*** and ***auditable***

# Similar Results with Other Privacy Decisions

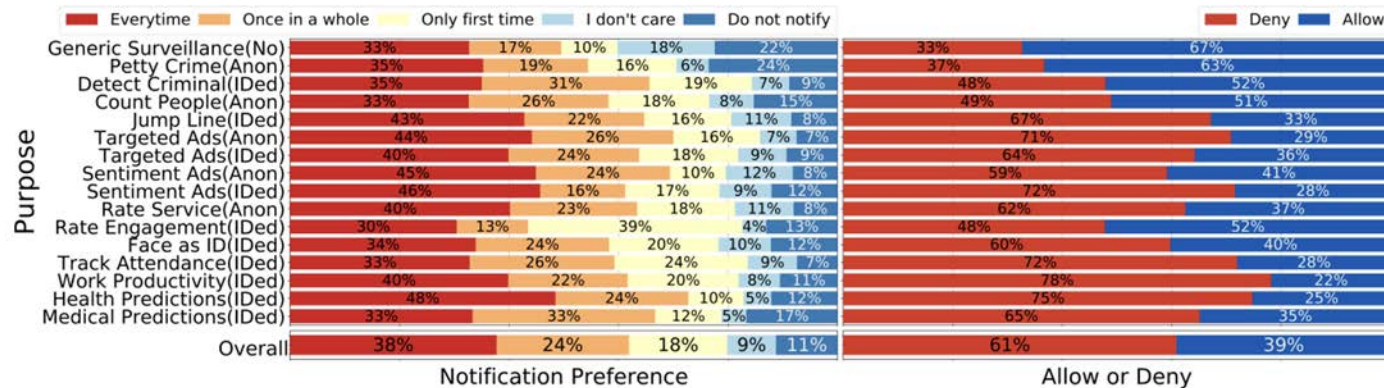
- Similar results with browser and IoT privacy decisions
- The challenge is that access to these privacy settings is generally not open
- **Would need regulation to make this possible...**
  - **...just think about the number of times you answer the same cookie questions...**

-S. Zhang, Y. Feng, A. Das, L. Bauer, L. Cranor, N. Sadeh, 'Understanding People's Privacy Attitudes Towards Video Analytics Technologies', CMU Sch. of Comp. Sci. Tech Report, CMU-ISR-20-114.  
-Daniel Smullen, Yuanyuan Feng, Shikun (Aerin) Zhang, Norman Sadeh, "The Best of Both Worlds: Mitigating Trade-offs Between Accuracy and User Burden in Capturing Mobile App Privacy Preferences", Privacy Enhancing Technologies Symposium (PETS 2020), Sep 2019 [\[pdf\]](#)

# In situ study of 123 people over 10 days in the context of their regular day-to-day activities



Deployment of a wide range of videoanalytics technologies:  
Very clear need for effective notification and control functionality



Very Clear Need for APIs and Privacy Assistants

Zhang, Y Feng, L Bauer, LF Cranor, A Das, and N Sadeh, "“Did you know this camera tracks your mood?”: Understanding Privacy Expectations and Preferences in the Age of Video Analytics", Proceedings on Privacy Enhancing Technologies, 2, 1, Apr 2021 [\[pdf\]](#)

# How about the Internet of Things?



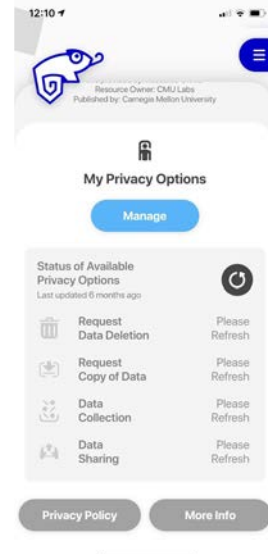
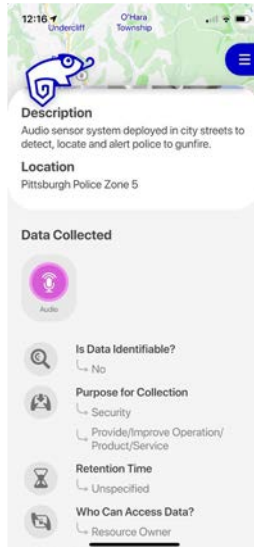
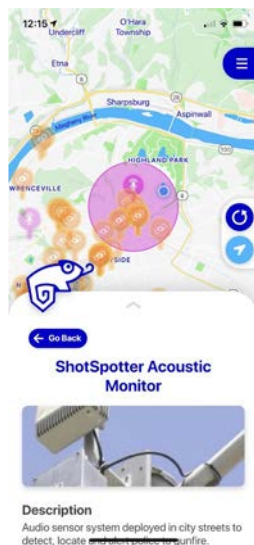
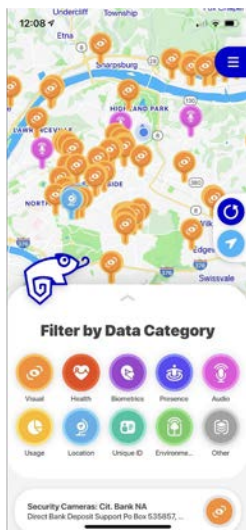
# IoT: Additional Usability Challenges



- How likely are you to notice this sign?
- Does this include facial recognition?
- What about facial expression or scene recognition?
- How long is the data retained?
- Do I get to opt in/opt out?
- Is this GDPR compliant?

# An IoT Privacy Infrastructure

- Portal to help people publicize the presence and data practices associated with IoT systems, incl. privacy controls (e.g. opt-in/out, deletion, etc.)
- IoT Assistant app (Android and iOS) for users to discover nearby IoT systems and their data practices, incl. accessing any available privacy controls



<https://www.iotprivacy.io>

US Patents 10,956,586 with additional USPTO and EPO patents pending

# IoT Privacy Infrastructure

- The IoT Assistant App [video](#)
- IoT Privacy Infrastructure Overview [video](#)
- Hosting nearly 150,000 IoT system entries today

A. Das, M. Degeling, D. Smullen, and N. Sadeh, Personalized Privacy Assistants for the Internet of Things, 2018 IEEE Pervasive Computing: Special Issue - Securing the IoT, April 2018

# Concluding Remarks - I

- Privacy is becoming **increasingly complex**
  - Everyone is collecting our data, increasingly complex data flows
  - Smartphones, IoT, AI/ML
- **New regulations** have been introduced that are in great part motivated by these developments (e.g., GDPR, CCPA/CPRA)
- These regulations are an important step in the right direction
- **Yet, in the absence of better technologies, they make usability even more challenging**
- AI is requiring people to take an increasingly broad view of privacy...Moving towards **a broader range of ethical considerations (e.g., from “access” to interpretability, explainability, “tweakability”)**

# Concluding Remarks - II

- My collaborators and I have been working on the development of technologies that aim to **mitigate these usability challenges**
- **Some successes over the years**
  - Introduction of increasingly finer permission settings in iOS and Android
  - Introduction of privacy labels in iOS and Android
  - Privacy nudges (e.g., Facebook, iOS)
  - Automated compliance tools for developers and regulators
  - Opt-Out Easy browser extension
  - Influences on public policy discussions/regulations (e.g. CCPA/CPRA, ADPPA)

# Concluding Remarks - II

- Our vision: **Privacy Assistants** that leverage
  - Techniques designed to empower people to take **advantage of more detailed privacy policy disclosures**
  - Techniques designed to motivate people and assist them with the **management** of an increasingly unmanageable number of **privacy decisions**
- Using techniques such as AI/ML or privacy nudges gives rise to complex **ethical issues** and requires **careful evaluation**
- Some of our efforts to help people manage their privacy settings call for **regulations that mandate privacy APIs - without these APIs privacy will remain unmanageable.**
- **Especially true in the Internet of Things**

# Q&A

The **Usable Privacy Policy Project** and the **Personalized Privacy Assistant Project** involve collaborations with a number of individuals

More details at:

***<https://usableprivacy.org>***

***<https://privacyassistant.org>***

***<https://explore.usableprivacy.org>***

***<https://www.iotprivacy.io>***